

# 9.1 Sampling

Students should be able to:

- identify a population and understand and explain the meaning of the terms 'parameter' and 'statistic'
- understand that information from a sample can be used to make inferences about the population it was taken from
- recognise and explain how to use sampling techniques, to include simple random sampling, systematic sampling, opportunity sampling, stratified sampling, quota sampling and cluster sampling
- understand that the use of prior information can make the sample more representative of the population (eg stratification)
- decide which sampling method to use to overcome practical problems when sampling
- discuss the advantages and disadvantages of each sampling method within a given context
- understand that increasing the sample size gives more reliable information about a population.

# **9.1 Sampling - Intro**

**Exercise 1.1 from  
HW10**

**Exercise 9.1A Q1 & 2  
Q1 summary &  
review**

# Exercise 1.1

- Q1**
- |             |             |
|-------------|-------------|
| a) Infinite | b) Finite   |
| c) Infinite | d) Infinite |
| e) Finite   | f) Finite   |

(6)

- Q2**
- a) The population is all the members of the book club. (1)
- b) A census should be used because all members of the book club should be consulted about the new book. Since it is a local book club, there should be few enough members to ask everyone. (2)

- Q3**
- a) A census should be done.  
The results will be more accurate and there are only 8 people in the population, so it wouldn't take long to find out the required information for each person. (2)
  - b) A sample survey should be done.  
Testing all 500 toys would take too long, but more importantly it would destroy all the toys. (2)
  - c) A sample survey is the only option.  
The population is all the possible dice rolls — there are an infinite number of dice rolls, so you can only examine a sample of them. (2)

**Q4** The sample is not random. For example: The ages of her mother's friends might not be representative of the entire population, or there might be a bias towards either men or women. Pooja's mother might have a group of friends who all have similar opinions on ethically-sourced products — this would introduce bias. Similarly, Pooja's mother might have very strong opinions on ethically-sourced products and give the questionnaires to people that she knows share her views.

(2)



# 9.1 Simple Random, Systematic

**Exercise 1.2 from  
HW10**

**Exercise 9.1A Q4**

## Exercise 1.2

**Q1** Use the random number generator to generate a list of 20 3-digit numbers, ignoring numbers outside of the range 001-500 and any repeated numbers. Then select the animals with the corresponding ID numbers. (3)

**Q2 a)** Not every student has an equal chance of being chosen. It's impossible to have two students from the same class and students in smaller classes have a greater chance of being picked. (1)

b) Number the students in the list from 001 to 700. Calculate the regular interval:  $700 \div 10 = 70$ . Generate a random number between 001 and 070 as the starting point, then keep adding 70 to this number. Select the students with the corresponding numbers for the sample. (4)

c) E.g. the systematic sample will include students from the whole age range, which should make it representative. The original method could have selected only younger or older students. (1)



Q3 a) The 108 dogs admitted to the sanctuary between 2015 and 2016. (1)

b) (i) Give each dog a unique 3-digit number between 001 and 108. Use a random number generator to generate 12 numbers from 001 to 108 (ignoring any that are larger than 108 or repeats) and select the dogs with the corresponding numbers. (2)

(ii) Give each dog a unique 3-digit number between 001 and 108. Calculate the regular interval:  $108 \div 12 = 9$ . Use a random-number generator to choose a starting point from 1 to 9. Keep adding 9 to the starting point and add all these dogs to the sample. (2)

**Q4** There are 365 (or 366) days in a year, so the regular interval is  $365 \div 50 = 7.3 \approx 7$  days. (1)

So, using this method, they will be collecting data once a week on the same day. Since the day of the week is likely to have an impact on the number of customers (e.g. there might be more customers on a Saturday), the sample is likely to be biased. (1)

# 9.1 Stratified, Quota

**Exercise 1.3 from  
HW10**

**Ex9.1A Q5, Q6**

## Exercise 1.3

Q1 Under 20: 3      20 to 40:  $2.2 \approx 2$

41 to 60: 1      Over 60:  $3.8 \approx 4$

(4)

Q2 Advantage: E.g. quota sampling is simple and convenient to carry out, as he doesn't need to search for specific athletes for the sample.

(1)

Disadvantage: E.g. depending on how Martin chooses the athletes, the sample may be biased, which could be avoided if a random sampling method was used instead.

(1)

Q3 a) Stratified sampling should be more representative of the population as it is likely that opinions will be different for the different areas (e.g. some areas may have better public transport than others). (1)

b) Addspatria: 35, Lessmorland: 43  
Conicston: 21, Logby Bridge: 34  
Angleside: 27, Gradesbeck: 40

(B1) for 2 correct → (3)

c) E.g. only the people who want to fill in the survey will send it back, which could mean that they get more responses from people with strong opinions. (1)



## **9.1 Opportunity, Cluster, Self-Selection**

**Exercise 1.4, 1.5,  
1.6 from HW10  
Exercise 9.1A Q3**

## Exercise 1.4

Q1 Advantage: E.g. it is simple and convenient to carry out. (1)

Disadvantage: E.g. the sample could be biased as it only includes the cars in one place at one time. (1)

Q2 E.g. The biologist cannot realistically identify and access the whole population, so opportunity sampling is practical. (1)

Q3 E.g. assuming that flipping the coin does not change or damage it in any way, the probability of getting heads would be the same in any flip, so the data won't be biased. (1)

**Q4** E.g. if the population is 'people in the UK', then his sample is likely to be very unrepresentative, since the members (1) are all in the same age range, and since they're friends they might like the same music. The sample could be improved by collecting data from other parts of the UK, other age ranges, etc. (1)

## Exercise 1.5

Q1 a) E.g. if the departments give different results to each other (e.g. because people in different departments have different working hours, or tend to live closer to/further from work), they would not make suitable clusters. (1)

b) Possible answers include:

Simple random sampling  
— avoids sampling bias.

Stratified sampling — Will can sample from different departments or age groups to reflect the population. (M1)

Quota sampling — convenient method that Will can use to sample from different departments or age groups to reflect the population. (A1)

**Q2** E.g. group the machines into clusters based on location, as you would expect machines in different locations to give similar results. Then randomly select enough clusters to give the required sample size and visit the companies in those clusters. Since they are travelling (E2) to the area anyway, it would make sense to do a one-stage cluster sample and collect data from every machine in the area. (1)



# Exercise 1.6

**Q1** E.g. she could advertise to the visitors (e.g. using a poster) to answer some questions, leaving her contact details (e.g. email address or phone number) for people to respond to. (1)

This is convenient as she does not need to identify the population of 'people who visit the park', and would save time compared to going to survey people in person. (1)

**Q2** Advantage: E.g. if there are differences in the reading habits of the boys and girls, making the sample more representative would prevent the higher number of responses from boys influencing the overall trends in the sample data. (1)

Disadvantage: E.g. if there is no overall difference in reading habits between boys and girls, then taking a sample of the data is unnecessary and is a waste of already collected data. (1)